C137 cont

These reagents can be directed towards a complementary antigen or antibody already attached to a VLSIPS™ Technology substrate.

Please replace the paragraph beginning on page 202, line 19, with the following rewritten paragraph:

C138

Positionally defined clones may be transferred to a new substrate by either physical transfer or by synthetic means. Synthetic means may involve either a production of the probe on the substrate using the VLSIPS™ Technology synthetic methods, or may involve the attachment of a targeting sequence made by VLSIPS™ Technology synthetic methods which will target that positionally defined clone to a position on a new substrate. Both methods will provide a substrate having a number of positionally defined probes useful in mapping.

REMARKS

Claims 172-209 are pending in this application and presented for examination. Attached is a marked up version of the changes made by the current Amendment. That attached page is captioned "VERSION WITH MARKINGS TO SHOW CHANGES MADE. The amendment to the specification corrects a number of informalities.

The nature of these changes include correction of typographical and grammatical and spelling errors, replacement or supplementation of serial numbers with U.S. Patent Numbers, insertion of ending page numbers for some references and insertion of trademarks when appropriate.

Entry of this amendment is respectfully urged since it merely cures formal defects in the specification of the application. Applicants submit that no new subject matter has been added.

Respectfully submitted,

Joseph R. Snyder
Reg. No. 39,381

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
Tel: (925) 472-5002
Fax: (415) 576-0300
JS/me
SF 1243411 v1

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In the Specification:

The paragraph beginning at page 1, line 1 has been amended as follows:

The present application is a continuation of 09/557,875, filed April 24, 2000, which is a continuation of 09/056,927, filed April 8, 1998, which is a continuation of 08/670,118 filed June 25, 1996, now U.S. Patent No. 5,800,992, which is a divisional of 08/168,904 filed December 15, 1993, which is a continuation of 07/624,114, filed December 6, 1990 (all incorporated by reference), which is a continuation-in-part of commonly assigned patent applications Pirrung et al., U.S.S.N. 07/362,901 (VLSIPS parent) filed on June 7, 1989; and Pirrung et al., U.S.S.N. 07/492,462 (VLSIPS CIP), filed on March 7, 1990 (now US 5,143,854), which are hereby incorporated herein by reference. The present application is also a continuation-in-part of [USSN 07/362,901, filed June 7, 1989. This application is also a continuation-in-part of] 08/348,471, filed November 30, 1994, which is a continuation of USSN 07/805,727, filed December 6, 1991 (now U.S. Patent No. 5,424,186), which is a continuation in part of USSN 07/624,120, filed December 6, 1990, which is a continuation-in-part of USSN 07/492,462, filed March 7, 1990 (now U.S. Patent No. 5,143,854), which is a continuation-in-part of USSN 07/362,901, filed June 7, 1989. Additional commonly assigned applications Barrett et al., U.S.S.N. 07/435,316 (caged biotin parent), filed November 13, 1989; and Barrett et al., U.S.S.N. 07/612,671 (caged biotin CIP), filed November 13, 1990 are also incorporated herein by reference. Additional applications Pirrung et al., U.S.S.N: 07/624,120 (now abandoned) a divisional of which has issued as US 5,744,101 and Dower et al., U.S.S.N. 07/626,730 (now US 5,547,839), which are also commonly assigned, are also hereby incorporated herein by reference.

The paragraph beginning at page 2, line 21 has been amended as follows:

Genetic information is critical in continuation of life processes. Life is substantially informationally based and its genetic content controls the growth and reproduction of the organism and its complements. Polypeptides, which are critical features of all living systems, are encoded by the genetic material of the cell. In particular, the properties of enzymes, functional proteins, and structural proteins are determined by the sequence of amino acids which make them up. As structure and function are integrally related, many biological functions may be explained by elucidating the underlying [the] structural features which provide those

functions. For this reason, it has become very important to determine the genetic sequences of nucleotides which encode the enzymes, structural proteins, and other effectors of biological functions. In addition to segments of nucleotides which encode polypeptides, there are many nucleotide sequences which are involved in control and regulation of gene expression.

The paragraph beginning at page 2, line 38 has been amended as follows:

The human genome project is directed toward determining the complete sequence of the genome of the human organism. Although such a sequence would not correspond to the sequence of any specific individual, it would provide significant information as to the general organization and specific sequences contained within segments from particular individuals. It would also provide mapping information which is very useful for further detailed studies. However, the need for highly rapid, accurate, and inexpensive sequencing technology is nowhere more apparent than in a demanding sequencing project such as this. To complete the sequencing of a human genome would require the determination of approximately $3x10^9$, or 3 billion base pairs.

The paragraph beginning on page 4, line 5 has been amended as follows:

The production of a substrate having a matrix of positionally defined regions with attached reagents exhibiting known recognition specificity can be used for the sequence analysis of a polymer. Although most directly applicable to sequencing, the present invention is also applicable to fingerprinting, mapping, and general screening of specific interactions. The VLSIPS™ Technology (Very Large Scale Immobilized Polymer Synthesis) substrates will be applied to evaluating other polymers, e.g., carbohydrates, polypeptides, hydrocarbon synthetic polymers, and the like. For these nonpolynucleotides, the sequence specific reagents will usually be antibodies specific for a particular subunit sequence.

The paragraph beginning on page 7, line 9, has been amended as follows:

A method of mapping a plurality of sequences relative to one another is also provided, the method comprising:

a) preparing a substrate having a plurality of positionally attached sequence specific probes [are] attached;

b) exposing each of said sequences to said substrate, thereby determining the patterns of interaction between said sequence specific probes and said sequences; and

c) determining the relative locations of said sequence specific probe interactions on said sequences to determine the overlaps and order of said sequences.

The paragraph beginning on page 8, line 15, has been amended as follows:

The detecting of the positions which bind the target sequence would typically be through a fluorescent label on the target. Although a fluorescent label is probably most convenient, other sorts of labels, e.g., radioactive, enzyme linked, optically detectable, or spectroscopic labels may be used. Because the oligonucleotide probes are positionally defined, the location of the hybridized duplex will directly translate to the sequences which hybridize. Thus, [upon] analysis of the positions provides a collection of subsequences found within the target sequence. These subsequences are matched with respect to their overlaps so as to assemble an intact target sequence.

The paragraph beginning on page 12, line 8, has been amended as follows:

Fig. 2 illustrates the proper function of a VLSIPS™ Technology [peptide] nucleotide synthesis.

The paragraph beginning on page 12, line 10, has been amended as follows:

Fig. 3 illustrates the proper function of a VLSIPS™ Technology [dipeptide] nucleotide synthesis.

The paragraph beginning on page 12, line 12, has been amended as follows:

Fig. 4 illustrates the process of a VLSIPS™ Technology trinucleotide synthesis.

The paragraph beginning on page 14, line 2, has been amended as follows:

I. Overall Description
   - A. general
   - B. VLSIPS™ Technology substrates
   - C. binary masking
   - D. applications
   - E. detection methods and apparatus

F. data analysis

The paragraph beginning on page 14, line 15, has been amended as follows:

III. Polynucleotide Sequencing
- A. preparation of substrate matrix
- B. labeling target polynucleotide
- C. hybridization conditions
- D. detection; VLSIPS™ Technology scanning
- E. analysis
- F. substrate reuse
- G. non-polynucleotide aspects

The paragraph beginning on page 14, line 24, has been amended as follows:

IV. Fingerprinting
- A. general
- B. preparation of substrate matrix
- C. labeling target nucleotides
- D. hybridization conditions
- E. detection; VLSIPS™ Technology scanning
- F. analysis
- G. substrate reuse
- H. non-polynucleotide aspects

The paragraph beginning on page 17, line 11, has been amended as follows:

B. VLSIPS™ Technology [Substrates]

The invention is enabled by the development of technology to prepare substrates on which specific reagents may be either positionally attached or synthesized. In particular, the very large scale- immobilized polymer synthesis [(]VLSIPS™ Technology [) technology] allows for the very high density production of an enormous diversity of reagents mapped out in a known matrix pattern on a substrate. These reagents specifically recognize subsequences in a target polymer and bind thereto, producing a map of positionally defined regions of interaction. These map positions are convertible into actual features recognized, and thus would be present in the target molecule of interest.

The paragraph beginning on page 17, line 28, has been amended as follows:

In the generic sense, the VLSIPS™ Technology [technology] allows the production of a substrate with a high density matrix of positionally mapped regions with specific recognition reagents attached at each distinct region. By use of protective groups which can be positionally removed, or added, the regions can be activated or deactivated for addition of

particular reagents or compounds. Details of the protection are described below and in related application [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854. In a preferred embodiment, photosensitive protecting agents will be used and the regions of activation or deactivation may be controlled by electro-optical and optical methods, similar to many of the processes used in semiconductor wafer and chip fabrication.

The paragraph beginning on page 18, line 3, has been amended as follows:

In the nucleic acid nucleotide sequencing application, a VLSIPS™ Technology substrate is synthesized having positionally defined oligonucleotide probes. See [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 ; and [U.S.S.N. ___/___,____] U.S. Patent No. 5,489,678[attorney docket number 11509-28 (automated VLSIPS)]. By use of masking technology and photosensitive synthetic subunits, the VLSIPS™ Technology apparatus allows for the stepwise synthesis of polymers according to a positionally defined matrix pattern. Each oligonucleotide probe will be synthesized at known -and defined positional locations on the substrate. This forms a matrix pattern of known relationship between position and specificity of interaction. The VLSIPS™ Technology [technology] allows the production of a very large number of different oligonucleotide probes to be simultaneously and automatically synthesized including numbers in excess of about 10^2, 10^3, 10^4, 10^5, 10^6, or even more, and at densities of at least about 10^2, $10^3/cm^2$, $10^4/cm^2$, $10^5/cm^2$ and up to $10^6/cm^2$ or more. This application discloses methods for synthesizing polymers on a silicon or other suitably derivatized substrate, methods and chemistry for synthesizing specific types of biological polymers on those substrates, apparatus for scanning and detecting whether interaction has occurred at specific locations on the substrate, and various other technologies related to the use of a high density very large scale immobilized polymer substrate. In particular, sequencing, fingerprinting, and mapping applications are discussed herein in detail, though related technologies are described in simultaneously filed applications U.S. Patent No. 5,489,678 [U.S.S.N. ___/____,____,attorney docket number 11509-28 (automated VLSIPS)] and [U.S.S.N.[____/____,____, 07/517,659] US Pat. No. 5,427,408 [attorney docket number 11509-16 (sequencing by synthesis)], each of which is hereby incorporated herein by reference.

The paragraph beginning on page 19, line 11, has been amended as follows:

These reagents could be individually attached at specific sites on the substrate in a matrix by an automated procedure where the regions are positionally targeted by some other

specific mechanism, e.g., one which would allow the entire collection of reagents to be attached to the substrate in a single reaction. Each reagent could be separately attached to a specific oligonucleotide sequence by an automated procedure. This would produce a collection of reagents where, e.g., each monoclonal antibody would have a unique oligonucleotide sequence attached to it. By virtue of a VLSIPS ™ Technology substrate which has different complementary oligonucleotides synthesized on it, each monoclonal antibody would specifically be bound only at that site on the substrate where the complementary oligonucleotide has been synthesized. A cross-linking step would fix the reagent to the substrate. See, e.g., Dattagupta et al. (1985) U.S. Patent No. 4,542,102 and (1987) U.S. Pat. No. 4,713,326; and Chatterjee, M. et al. (1990) J. Am. Chem. Soc. 112:[6997-_____] 6397-6399, which are hereby incorporated herein by reference. This allows a high density positionally specific collection of specific recognition reagents, e.g., monoclonal antibodies, to be immobilized to a solid substrate using an automated system.

The paragraph beginning on page 82, line 2, has been amended as follows:

In fact, the means for producing a substrate useful for these techniques are explained in [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854, which is hereby incorporated herein by reference. However, there are various particular ways to optimize the synthetic processes. Many of these methods are described in [U.S.S.N. ____/_____,_____] U.S. Patent No. 5,489,678[, attorney docket number 11509-28 (automated VLSIPS)].

The paragraph beginning on page 86, line 2, has been amended as follows:

VLSIPS™ **TECHNOLOGY** PROJECT IN NUCLEIC ACIDS

The paragraph beginning on page 87, line 8, has been amended as follows:

In addition to the sequencing, fingerprinting, and mapping applications, the present invention also provides means for determining specificity of interaction with particular sequences. Many of these applications were described in U.S.S.N. 07/362,901 [(VLSIPS parent)], from which CIP U.S.S.N. 07/492,462[,] issued as [(VLSIPS CIP)] U.S. Pat. No. 5,143,854, and U.S.S.N. 07/435,316 [(caged biotin parent), and], from which CIP U.S.S.N. 07/612,671 [(caged biotin CIP)] issued as U.S. Pat. No. 5,252,743.

The paragraph beginning on page 87, line 34, has been amended as follows:

With a detection method selected, an apparatus for scanning the substrate will be designed. Apparatus, as described in U.S.S.N. 07/362,901, [(VLSIPS parent); or] from which CIP U.S.S.N. 07/492,462 [(VLSIPS CIP)] issued as U.S. Pat. No. 5,143,854; or [U.S.S.N. ____/____,____] U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (automated VLSIPS),] are particularly appropriate. Design modifications may also be incorporated therein.

The paragraph beginning on page 88, line 5, has been amended as follows:

Data is analyzed by processes similar to those described below in the section describing theoretical analysis. More efficient algorithms will be mathematically devised, and will usually be designed to be performed on a computer. Various computer programs which may more quickly or efficiently make measurement samples and distinguish signal from noise will also be devised. See, particularly, U.S. Patent No. 5,489,678. [U.S.S.N. ____/___,____, attorney docket number 11509-28 (automated VLSIPS).]

The paragraph beginning on page 88, line 37, has been amended as follows:

Accordingly, in one embodiment of the invention the data are corrected for removal of these spurious data points, and an average of the data points is thereafter utilized in determining relative binding efficiency. In general the data are fitted to a base curve and [statistically] measures are used to remove spurious data.

The paragraph beginning on page 89, line 12, has been amended as follows:

The principle of the hybridization sequencing procedure is based, in part, upon the ability to determine overlaps of short segments. The VLSIPS™ Technology [technology] provides the ability to generate reagents which will saturate the possible short subsequence recognition possibilities. The principle is most easily illustrated by using a binary sequence, such as a sequence of zeros and ones. Once having illustrated the application to a binary alphabet, the principle may easily be understood to encompass three letter, four letter, five or more. letter, even 20 letter alphabets. A theoretical treatment of analysis of subsequence information, to reconstruction of a target sequence is provided, [e.e.] e.g., in Lysov, Yu., et al. (1988) Doklady Akademi. Nauk. SSR 303:1508-1511; [Khropko] Khrapko K., et al. (1989) FEBS Letters 256:118-122; Pevzner, P. (1989) J. of Biomolecular Structure and Dynamics 7:63-69; and

Drmanac, R. et al. (1989) Genomics 4:114-128; each of which is hereby incorporated herein by reference.

The paragraph beginning on page 90, line 7, has been amended as follows:

A simple example is presented below of how a sequence of ten digits comprising zeros and ones would be sequenceable using short segments of five digits. For example, consider the sample ten digit sequence:

1010011100.

A VLSIPS™ Technology substrate could be constructed, as discussed elsewhere, which would have reagents attached in a defined matrix pattern which specifically recognize each of the possible five digit sequences of ones and zeros. The number of possible five digit subsequences is $2^5 = 32$. The number of possible different sequences 10 digits long is $2^{10} = 1,024$. The five contiguous digit subsequences within a ten digit sequence number six, i.e., positioned at digits 1-5, 2-6, 3-7, 4-8, 5-9, and 6-10. It will be noted that the specific order of the digits in the sequence is important and that the order is directional, e.g., running left to right versus right to left. The first five digit sequence contained in the target sequence is 10100. The second is 01001, the third is 10011, the fourth is 00111, the fifth is 01110, and the sixth is 11100.

The paragraph beginning on page 90, line 27, has been amended as follows:

The VLSIPS™ Technology substrate would have a matrix pattern of positionally attached reagents which recognize each of the different 5-mer subsequences. Those reagents which recognize each of the 6 contained 5-mers will bind the target, and a label allows the positional determination of where the sequence specific interaction has occurred. By correlation of the position in the matrix pattern, the corresponding bound subsequences can be determined.

The paragraph beginning on page 91, line 24, has been amended as follows:

In this example, not only do we know that the sequence contains 10100, but we also know that it contains the second five character sequence, 01001. By virtue of knowing that the sequence contains 10100, we can look specifically to determine whether the sequence contains a subsequence of five characters which contains the four leftmost digits plus a next digit to the left. For example, we would look for a sequence of X1010, but we find that there is none. Thus, we know that the 10100 must be at the left end of the 10-mer. We would also look to see whether the sequence contains the rightmost four digits-plus a next digit to the right, e.g., 0100X.

We find that the sequence also contains the sequence 01001, and that X is a 1. Thus, we know at least that our target sequence has an overlap of 0100 and has the left terminal sequence 101001.

The paragraph beginning on page 93, line 28, has been amended as follows:

A four letter alphabet may be conceptualized in at least two different ways from the two letter alphabet. One way[,] is to consider the four possible values at each position and to analogize in a similar fashion to the binary example each of the overlaps. A second way is to group the binary digits into groups.

The paragraph beginning on page 98, line 23, has been amended as follows:

By way of further explanation, all possible oligonucleotide 8-mers may be depicted in the fashion:

N1-N2-N3-N4-N5-N6-N7-N8,

in which there are $4^8 = 65,536$ possible 8-mers. As described in [U.S.S.N. __/___,___] U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (]automated VLSIPS™ Technology[)], producing all possible 8-mers requires 4x8 = 32 chemical binary synthesis steps to produce the entire matrix pattern of 65,536 8-mer possibilities. By incorporating degeneracy reducing nucleotides, D's, which hybridize nonselectively to any corresponding complementary nucleotide,

new oligonucleotides 12-mers can be made in the fashion:

N1-N2-N3-N4-D-D-D-D-N5-N6-N7-N8,

in which there are again, as above, only $4^8 = 65,536$ possible "12-mers", which in reality only have 8 different nucleotides.

The paragraph beginning on page 99, line 31, has been amended as follows:

For other polymer targets, the specific reagents will often be polypeptides. These polypeptides may be protein binding domains from enzymes or other proteins which display specificity for binding. Usually an antibody molecule may be used, and monoclonal antibodies may be particularly desired. Classical methods may be applied for preparing antibodies, see, e.g., Harlow and Lane (1988) Antibodies: A Laboratory Manual Cold Spring Harbor Press, New York; and Goding (1986) Monoclonal Antibodies: Principles and Practice (2d Ed.) Academic Press, San Diego. Other suitable techniques for in vitro exposure of lymphocytes to the antigens

or selection of libraries of antibody binding sites are described, e.g., in Huse et al. (1989) Science 246:1275-1281; and Ward et al. 91989) Nature 341:544-546, each of which is hereby incorporated herein by reference. Unusual antibody production methods are also described, e.g., in Hendricks et al. (1989) BioTechnology 7:1271-1274; and Hiatt et al. (1989) Nature 342:76-78, each of which is hereby incorporated herein by reference. Other molecules which may exhibit specific binding interaction may be useful for attachment to a VLSIPS™ technologysubstrate by various methods, including the caged biotin methods, see, e.g., [U.S.S.N. 07/435,316, (caged biotin parent) U.S. Patent No. 5,252,743, and U.S.S.N. 07/612,671 (caged biotin CIP)] Barrett et al. (1993) U.S. Pat. No. 5,252,743.

The paragraph beginning on page 100, line 16, has been amended as follows:

The antibody specific reagents should be particularly useful for the polypeptide, carbohydrate, and synthetic polymer applications. Individual specific reagents might be generated by an automated process to generate the number of reagents necessary to advantageously use the high density positional matrix pattern. In an alternative approach, a plurality of hybridoma cells may be screened for their ability to bind to a VLSIPS™ Technology matrix possessing the desired sequences whose binding specificity is desired. Each cell might be individually grown up and its binding specificity determined by VLSIPS™ Technology apparatus and technology. An alternative strategy would be to expose the same VLSIPS™ Technology matrix to a polyclonal serum of high titer. By a successively large volume of serum and different animals, each region of the VLSIPS™ Technology substrate would have attached to it a substantial number of antibody molecules with specificity of binding. The substrate, with non-covalently bound antibodies could be derivatized and the antibodies transferred to an adjacent second substrate in the matrix pattern in which the antibody molecules had attached to the first matrix. If the sensitivity of detection of binding interaction is sufficiently high, such a low efficiency transfer of antibody molecules may produce a sufficiently high signal to be useful for many purposes, including the sequencing applications.

The paragraph beginning on page 101, line 10, has been amended as follows:

In principle, the making of a substrate having a positionally defined matrix pattern of all possible oligonucleotides of a given length involves a conceptually simple method of synthesizing each and every different possible oligonucleotide, and [affixed] affixing them to a definable position. Oligonucleotide synthesis is presently mechanized and enabled by current

technology, see, e.g., [U.S.S.N. 07/362,901, (VLSIPS parent); U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854; and instruments supplied by Applied Biosystems, Foster City, California.

The paragraph beginning on page 101, line 34, has been amended as follows:

Once the desired repertoire of possible oligomer sequences of a given length have been synthesized, this collection of reagents may be individually positionally attached to a substrate, thereby allowing a batchwise hybridization step. Present technology also would allow the possibility of attaching each and every one of these 10-mers to a separate specific position on a solid matrix. This attachment could be automated in any of a number of ways, particularly through the use of a caged biotin type linking. This would produce a matrix having each of different possible 10-mers.

The paragraph beginning on page 102, line 6, has been amended as follows:

A batchwise hybridization is much preferred because of its reproducibility and simplicity. An automated process of attaching various reagents to positionally defined sites on a substrate is provided in [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854; [U.S.S.N. __/___,___, attorney docket number 11509-28 (automated VLSIPS)] U.S. Pat. No. 5,489,678; and [U.S.S.N. 07/612,671 (caged biotin CIP)] Barrett et al. (1993) U.S. Pat. No. 5,252,743; each of which is hereby incorporated herein by reference.

The paragraph beginning on page 102, line 13, has been amended as follows:

Instead of separate synthesis of each oligonucleotide, these oligonucleotides are conveniently synthesized in parallel by sequential synthetic processes on a defined matrix pattern as provided in [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854; and [U.S.S.N. __/___,___] U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (automated VLSIPS),] which are incorporated herein by reference. Here, the oligonucleotides are synthesized stepwise on a substrate at positionally separate and defined positions. Use of photosensitive blocking reagents allows for defined sequences of synthetic steps over the surface of a matrix pattern. By use of the binary masking strategy, the surface of the substrate can be positioned to generate a desired pattern of regions, each having a defined sequence oligonucleotide synthesized and immobilized thereto.

The paragraph beginning on page 102, line 27, has been amended as follows:

Although the prior art technology can be used to generate the desired repertoire of oligonucleotide probes, an efficient and cost effective means would be to use the VLSIPS technology described in[U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 and [U.S.S.N. __/___,___,] [attorney docket number 11509-28 (automated VLSIPS)]U.S. Patent No. 5,489,678. In this embodiment, the photosensitive reagents involved in the production of such a matrix are described below.

The paragraph beginning on page 103, line 4, has been amended as follows:

At a size of about 30 microns by 30 microns, one million regions would take about 11 centimeters square or a single wafer of about 4 centimeters by 4 centimeters. Thus the present technology provides for making a single matrix of that size having all one million plus possible oligonucleotides. Region size [are] is sufficiently small to correspond to densities of at least about 5 regions/cm^2, 20 regions/cm^2, 50 regions/cm^2, 100 regions/cm^2, and greater, including 300 regions/cm^2, 1000 regions/cm^2, 3K regions/cm^2, 10K regions/cm^2, 30K regions/cm^2, 100K regions/cm^2, 300K regions/cm^2 or more, even in excess of one million regions/cm^2.

The paragraph beginning on page 103, line 15, has been amended as follows:

Although the pattern of the regions which contain specific sequences is theoretically not important, for practical reasons certain patterns will be preferred in synthesizing the oligonucleotides. The application of binary masking algorithms for generating the pattern of known oligonucleotide probes is described in related [U.S.S.N. __/___,___] U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (automated VLSIPS) which was filed simultaneously with this application]. By use of these binary masks, a highly efficient means is provided for producing the substrate with the desired matrix pattern of different sequences. Although the binary masking strategy allows for the synthesis of all lengths of polymers, the strategy may be easily modified to provide only polymers of a given length. This is achieved by omitting steps where a subunit is not attached.

The paragraph beginning on page 103, line 30, has been amended as follows:

The strategy for generating a specific pattern may take any of a number of different approaches. These approaches are well described in related application [U.S.S.N.

__/___,___]U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (automated VLSIPS)] and include a number of binary masking approaches which will not be exhaustively discussed herein. However, the binary masking and binary synthesis approaches provide a maximum of diversity with a minimum number of actual synthetic steps.

The paragraph beginning on page 104, line 1, has been amended as follows:

The length of oligonucleotides used in sequencing applications will be selected on criteria determined to some extent by the practical limits discussed above. For example, if probes are made as oligonucleotides, there will be 65,536 possible eight nucleotide sequences. If a nine subunit oligonucleotide is selected, there are 262,144 possible permeations of sequences. If a ten-mer oligonucleotide is selected, there are 1,048,576 possible permeations of sequences. As the number gets larger, the required number of positionally defined subunits necessary to saturate the possibilities also increases. With respect to hybridization conditions, the length of the matching necessary to [converse] confer stability of the conditions selected can be compensated for. See, e.g., Kanehisa, M. (1984) Nuc. Acids Res. 12:203-213, which is hereby incorporated herein by reference.

The paragraph beginning on page 104, line 16, has been amended as follows:

Although not described in detail here, but below for oligonucleotide probes, the VLSIPS™ Technology [technology] would typically use a photosensitive protective group on an oligonucleotide. Sample oligonucleotides are shown in Figure 1. In particular, the photoprotective group on the nucleotide molecules may be selected from a wide variety of positive light reactive groups preferably including nitro aromatic compounds such aso-nitro-benzyl derivatives or benzylsulfonyl. See, e.g., Gait (1984) Oligonucleotide Synthesis: A Practical Approach, IRL Press, Oxford, which is hereby incorporated herein by reference. In a preferred embodiment, 6-nitro-veratryl oxycarbony (NVOC), 2-nitrobenzyl oxycarbonyl (NBOC), or α,α-dimethyl-dimethoxybenzyl oxycarbonyl (DEZ) is used. Photoremovable protective groups are described in, e.g., Patchornik (1970) J. Amer. Chem. Soc. 92:6333-[____] 6335; and Amit et al. (1974) J. Organic Chem. 39:192-[___] 196; each of which is hereby incorporated herein by reference.

The paragraph beginning on page 105, line 13, has been amended as follows:

Then, the target may be bound to the whole collection of beads and those beads that have appropriate specific reagents on them will bind to the target. Then a sorting system may be utilized to sort those beads that actually bind the target from those that do not. This may be accomplished by presently available cell sorting devices or a similar apparatus. After the relatively small number of beads which have bound the target have been collected, the encoding scheme may be read off to determine the specificity of the reagent on the bead. An encoding system may include a magnetic system, a shape encoding system, a color encoding system, or a combination of any of these, or any other encoding system. Once again, with the collection of specific interactions that have occurred, the binding may be analyzed for sequence information, fingerprint information, or mapping information.

The paragraph beginning on page 106, line 13, has been amended as follows:

The length of the probe is selected for a length that[it] will allow the probe to bind with specificity to possible targets. The hybridization conditions are also very important in that they will determine how [close] closely the homology of complementary binding will be detected. In fact, a single target may be evaluated at a number of different conditions to determine its spectrum of specificity for binding particular probes. This may find use in a number of other applications besides the polynucleotide sequencing fingerprinting or mapping. For example, it will be desired to determine the spectrum of binding affinities and specificities of cell surface antigens with binding by particular antibodies immobilized on the substrate surface, particularly under different interaction conditions. In a related fashion, different regions with reagents having differing affinities or levels of specificity may allow such a spectrum to be defined using a single incubation, where various regions, at a given hybridization condition, show the binding affinity. For example, fingerprint probes of various lengths, or with specific defined non-matches may be used. Unnatural nucleotides or nucleotides exhibiting modified specificity of complementary binding are described in greater detail in Macevicz (1990) PCT pub. No. WO 90/04652; and see the section on modified nucleotides in the Sigma Chemical Company catalogue.

The paragraph beginning on page 108, line 20, has been amended as follows:

The hybridization conditions between probe and target should be selected such that the specific recognition interaction, i.e., hybridization, of the two molecules is both sufficiently specific and sufficiently stable. See, e.g., Hames and Higgins (1985) Nucleic Acid

Hybridisation: A Practical Approach, IRL Press, Oxford. These conditions will be dependent both on the specific sequence and often on the guanine and cytosine (GC) content of the complementary hybrid strands. The conditions may often be selected to be universally equally stable independent of the specific sequences involved. This typically will make use of a reagent such as an [arylammonium] alkylammonium buffer. See, Wood et al. (1985) "Base Composition-independent Hybridization in Tetramethylammonium Chloride: A Method for Oligonucleotide Screening of Highly Complex Gene Libraries," Proc. Natl. Acad. Sci. USA, 82:1585-1588; and Krupov et al. (1989) "An Oligonucleotide Hybridization Approach to DNA Sequencing," FEBS Letters, 256:118-122; each of which is hereby incorporated herein by reference. An [arylammonium] alkylammonium buffer tends to minimize differences in hybridization rate and stability due to GC content. By virtue of the fact that sequences then hybridize with approximately equal affinity and stability, there is relatively little bias in strength or kinetics of binding for particular sequences. Temperature and salt conditions along with other buffer parameters should be selected such that the kinetics of renaturation should be essentially independent of the specific target subsequence or oligonucleotide probe involved. In order to ensure this, the hybridization reactions will usually be performed in a single incubation of all the substrate matrices together exposed to the identical same target probe solution under the same conditions.

The paragraph beginning on page 109, line 13, has been amended as follows:

Alternatively, various substrates may be individually treated differently. Different substrates may be produced, each having reagents which bind to target subsequences with substantially identical stabilities and kinetics of hybridization. For example, all of the high GC content probes could be synthesized on a single substrate which is treated accordingly. In this embodiment, the [arylammonium] alkylammonium buffers could be unnecessary. Each substrate is then treated in a manner such that the collection of substrates show essentially uniform binding and the hybridization data of target binding to the individual substrate matrix is combined with the data from other substrates to derive the necessary subsequence binding information. The hybridization conditions will usually be selected to be sufficiently specific such that the fidelity of base matching will be properly discriminated. Of course, control hybridizations should be included to determine the stringency and kinetics of hybridization.

The paragraph beginning on page 109, line 31, has been amended as follows:

D. Detection; VLSIPS™ **Technology** Scanning

The next step of the sequencing process by hybridization involves labeling of target polynucleotide molecules. A quickly and easily detectable signal is preferred. The VLSIPS™ Technology apparatus is designed to easily detect a fluorescent label, so fluorescent tagging of the target sequence is preferred. Other suitable labels include heavy metal labels, magnetic probes, chromogenic labels (e.g., phosphorescent labels, dyes, and fluorophores) spectroscopic labels, enzyme linked labels, radioactive labels, and labeled binding proteins. Additional labels are described in U.S. Pat. No. 4,366,241, which is incorporated herein by reference.

The paragraph beginning on page 110, line 5, has been amended as follows:

The detection methods used to determine where hybridization has taken place will typically depend upon the label selected above. Thus, for a fluorescent label a fluorescent detection step will typically be used. [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 and [U.S.S.N. __/___,___] U.S. Patent No. 5,489,678[, attorney docket number 11509-28 (automated VLSIPS)] describe apparatus and mechanisms for scanning a substrate matrix using fluorescence detection, but a similar apparatus is adaptable for other optically detectable labels.

The paragraph beginning on page 112, line 3, has been amended as follows:

Typically, an antigen, or collection of antigens are presented to an immune system. This may take the form of synthesized short polymers produced by the VLSIPS™ Technology [technology], or by the other synthetic means, or from isolation of natural products. For example, antigen for the polypeptides may be made by the VLSIPS™ Technology [technology], by standard peptide synthesis, by isolation of natural proteins with or without degradation to shorter segments, or by expression of a collection of short nucleic acids of random or defined sequences. See, e.g., Tuerk and Gold (1990) Science 249:505-510, for generation of a collection of randomly mutagenized oligonucleotides useful for expression.

The paragraph beginning on page 112, line 21, has been amended as follows:

A large diversity of antibodies will be generated, some of which have specificities for the desired sequences. Antibodies may be purified having the desired sequence specificities by isolating the cells producing them. For example, a VLSIPS™ Technology substrate with the

desired antigens synthesized thereon may be used to isolate cells with cell surface reagents which recognize the antigens. The VLSIPS™ Technology substrate may be used as an affinity reagent to select and recover the appropriate cells. Antibodies from those cells may be attached to a substrate using the caged biotin methodology, or by attaching a targeting molecule, e.g., an oligonucleotide. Alternatively, the supernatants from antibody producing cells can be easily assayed using a VLSIPS™ Technology substrate to identify the cells producing the appropriate antibodies.

The paragraph beginning on page 113, line 7, has been amended as follows:

In one particular embodiment, a VLSIPS™ Technology substrate, e.g., with a large plurality of fingerprint antigens attached thereto, is used to isolate antibodies from a supernatant of a population of cells producing antibodies to the antigens. Using the substrate as an affinity reagent, the antibodies will attach to the appropriate positionally defined antigens. The antibodies may be carefully removed therefrom, preferably by an automated system which retains their homogeneous specificities. The isolated antibodies can be attached to a new substrate in a positionally defined matrix pattern.

The paragraph beginning on page 113, line 17, has been amended as follows:

In a further embodiment, these spatially separated antibodies may be isolated using a specific targeting method for isolation. In this embodiment, a linker molecule which attaches to a particular portion of the antibody, preferably away from the binding site, can be attached to the antibodies. Various reagents will be used, including staphylococcus protein A or antibodies which bind to domains remote from the binding site. Alternatively, the antibodies in the population, before affinity purification, may be derivatized with an appropriate reagent compatible with new VLSIPS™ Technology synthesis. A preferred reagent is a nucleotide which can serve as a linker to synthetic VLSIPS™ Technology steps for synthesizing a specific sequence thereon. Then, by successive VLSIPS™ Technology cycles, each of the antibodies attached to the defined antigen regions can have a defined oligonucleotide synthesized thereon and corresponding in area to the region of the substrate having each antigen attached. These defined oligonucleotides will be useful as targeting reagents to attach those antibodies possessing the same target sequence specificity at defined positions on a new substrate, by virtue of having bound to the antigen region, to a new VLSIPS™ Technology substrate having the complementary target oligonucleotides positionally located on it. In this fashion, a VLSIPS™ Technology

substrate having the desired antigens attached thereto can be used to generate a second VLSIPS™ Technology substrate with positionally defined reagents which recognize those antigens.

The paragraph beginning on page 114, line 33, has been amended as follows:

A collection of specific probes may be produced by either of the methods described above in the section on sequencing. Specific oligonucleotide probes of desired lengths may be individually synthesized on a standard oligonucleotide synthesizer. The length of these probes is limited only by [the length of] the ability of the synthesizer to continue to accurately synthesize a molecule. Oligonucleotides or sequence fragments may also be isolated from natural sources. Biological amplification methods may be coupled with synthetic synthesizing procedures such as, e.g., polymerase chain reaction.

The paragraph beginning on page 115, line 6, has been amended as follows:

In one embodiment, the individually isolated probes may be attached to the matrix at defined positions. These probe reagents may be attached by an automated process making use of the caged biotin methodology described in U.S. Patent No. 5,252,743 [S.N. 07/612,671 (caged biotin CIP)], or using photochemical reagents, see, e.g., Dattagupta et al. (1985) U.S. Pat. No. 4,542,102 and (1987) U.S. Pat. No. 4,713,326. Each [individual] individually purified reagent can be attached individually at specific locations on a substrate.

The paragraph beginning on page 115, line 15, has been amended as follows:

In another embodiment, the VLSIPS™ Technology synthesizing technique may be used to synthesize the desired probes at specific positions on a substrate. The probes may be synthesized by successively adding appropriate monomer subunits, e.g., nucleotides, to generate the desired sequences.

The paragraph beginning on page 115, line 20, has been amended as follows:

In another embodiment, a relatively short specific oligonucleotide is used which serves as a targeting reagent for positionally directing the sequence recognition reagent. For example, the sequence specific reagents having a separate additional sequence recognition segment (usually of a different polymer from the target sequence) can be directed to target oligonucleotides attached to the substrate. By use of non-natural targeting reagents, e.g., unusual nucleotide analogues which pair with other unnatural nucleotide analogues and which do not

interfere with natural nucleotide interactions, the natural and non-natural portions can coexist on the same molecule without interfering with their individual functionalities. This can combine both a synthetic and biological production system analogous to the technique for targeting monoclonal antibodies to locations on a VLSIPS™ Technology substrate at defined positions. Unnatural optical isomers of nucleotides may be useful unnatural reagents subject to similar chemistry, but incapable of interfering with the natural biological polymers. See also, U.S Patent No. 5,547,839[U.S.S.N. __/___,___} [attorney docket number 11509-26 (sequencing by synthesis);] which is hereby incorporated herein by reference.

The paragraph beginning on page 116, line 3, has been amended as follows:

After the separate substrate attached reagents are attached to the targeting segment, the two are crosslinked, thereby permanently attaching them to the substrate. Suitable crosslinking reagents are known, see, e.g., Dattagupta et al. (1985) U.S. Pat. No. 4,542,102 and (1987) "Coupling of nucleic acids to solid support by photochemical methods," U.S. Pat. No. 4,713,326, each of which is hereby incorporated herein by reference. Similar linkages for attachment of proteins to a solid substrate are provided, e.g., in Merrifield (1986) Science 232:341-[____]347, which is hereby incorporated herein by reference.

The paragraph beginning on page 117, line 26, has been amended as follows:

E. Detection; VLSIPS™ **Technology** Scanning

The paragraph beginning on page 118, line 4, has been amended as follows:

Of course, the VLSIPS™ Technology scanning apparatus may also be useful to generate a digitized version of the fingerprint pattern. In this way, the identification pattern can be provided in a linear string of digits. This sequence could also be used for a standardized identification system providing significant useful medical transferability of specific data. In one embodiment, the probes used are selected to be of sufficiently high resolution to measure the antigens of the major histo compatibility complex[,]. [it] It might even be possible to provide transplantation matching data in a linear stream of data. The fingerprinting data may provide a condensed version, or summary, of the linear genetic data, or any other information data base.

The paragraph beginning on page 118, line 36, has been amended as follows:

Besides polynucleotide applications, the fingerprinting analysis may be applied to other polymers, especially polypeptides, carbohydrates, and other polymers, both organic and inorganic. Besides using the fingerprinting method for analyzing a particular polymer, the fingerprinting method may be used to characterize various samples. For example, a cell or population of cells may be tested for their expression of specific antigens or their mRNA sequence intent. For example, a T-cell may be classified by virtue of its combination of expressed surface antigens. With specific reagents which interact with these antigens, a cell or a population of cells or a lysed cell may be exposed to a VLSIPS™ Technology substrate. The biological sample may be classified or characterized by analyzing the pattern of specific interaction. This may be applicable to a cell or tissue type, to the[expressed] messenger RNA population expressed by a cell to the genetic content of a cell, or to virtually any sample which can be classified and/or identified by its combination of specific molecular properties.

The paragraph beginning on page 120, line 31, has been amended as follows:

The high resolution VLSIPS™ Technology substrate may also be used as a very powerful diagnostic tool to test the combination of presence, of a plurality of different assays from a biological sample. For example, a cancerous condition may be indicated by a combination of various different properties found in the blood. For example, a cancerous condition may be indicated by a combination of expression of various soluble antigens found in the blood along with a high number of various cellular antigens found on lymphocytes and/or particular cell degradation products. With a substrate as provided herein, a large number of different features can be simultaneously performed on a biological sample. In fact, the high resolution of the test will allow more complete characterization of parameters which define particular diseases. Thus, the power of diagnostic tests may be limited by the extent of statistical correlation with a particular condition rather than with the number of antigens or interactions which are tested. The present invention provides the means to generate this large universe of possible reagents and the ability to actually accumulate that correlative data.

The paragraph beginning on page 121, line 13, has been amended as follows:

In another embodiment, a substrate as provided herein may be used for genetic screening. This would allow for simultaneous screening of thousands of genetic markers. As the density of the matrix is increased, many more molecules can be simultaneously tested. Genetic screening then becomes a simpler method as the present invention provides the ability to screen

for thousands, tens of thousands, and hundreds of thousands, even millions of different possible genetic features. However, the number of high correlation genetic markers for conditions numbers only in the hundreds. Again, the possibility for screening a large number of sequences provides the opportunity for generating the data which can provide correlation between sequences and specific conditions or susceptibility. The present invention provides the means to generate extremely valuable correlations useful for the genetic detection of the causative mutation leading to medical conditions. In still another embodiment, the present invention would be applicable to distinguishing two individuals having identical genetic compositions. The antibody population within an individual is dependent both on genetic and historical factors. Each individual experiences a unique exposure to various infectious agents, and the combined antibody expression is partly determined thereby. Thus, individuals may also be fingerprinted by their immunological content, either of actively expressed antibodies, or their immunological memory. Similar sorts of immunological and environmental histories may be useful for fingerprinting, perhaps in combination with other screening properties. In particular, the present invention may be useful for screening allergic reactions or susceptibilities, and a simple IgE specificity test may be useful in determining a spectrum of allergies.

The paragraph beginning on page 123, line 7, has been amended as follows:

With the [fingerprinted] fingerprinting method as [in] an identification means arises from [mosaism] mosaicism problems in an organism. A mosaic organism is one whose genetic content in different cells is significantly different. Various clonal populations should have similar genetic fingerprints, though different clonal populations may have different genetic contents. See, for example, Suzuki et al. An Introduction to Genetic Analysis (4th Ed.), Freeman and Co., New York, which is hereby incorporated herein by reference. However, this problem should be a relatively rare problem and could be more carefully evaluated with greater experience using the fingerprinting methods.

The paragraph beginning on page 125, line 2, has been amended as follows:

The labeling methods will be similar to those applicable in sequencing and fingerprinting embodiments. [Again, the target sequences may be desired to be fragmented.] Again, it may be desirable to fragment the target sequences.

The paragraph beginning on page 127, line 18, has been amended as follows:

As originally indicated in the parent filing of VLSIPS™ Technology, the production of a high density plurality of spatially segregated polymers provides the ability to generate a very large universe or repertoire of individually and distinct sequence possibilities. As indicated above, particular oligonucleotides may be synthesized in automated fashion at specific locations on a matrix. In fact, these oligonucleotides may be used to direct other molecules to specific locations by linking specific oligonucleotides to other reagents which are in batch exposed to the matrix and hybridized in a complementary fashion to only those locations where the complementary oligonucleotide has been synthesized on the matrix. This allows for spatially attaching a plurality of different reagents onto the matrix instead of individually attaching each separate reagent at each specific location. Although the caged biotin method allows[the] automated attachment, the speed of the caged biotin attachment process is relatively slow and requires a separate reaction for each reagent being attached. By use of the oligonucleotide method, the specificity of position can be done in an automated and parallel fashion. As each reagent is produced, instead of directly attaching each reagent at each desired position, the reagent may be attached to a specific desired complementary oligonucleotide which will ultimately be specifically directed toward locations on the matrix having a complementary oligonucleotide attached thereat.

The paragraph beginning on page 129, line 2, has been amended as follows:

Once a gene is sequenced, the present invention provides a means to compare alleles or related sequences to locate and identify differences from the control sequence. This would be extremely useful in further analysis of genetic variability at a specific gene locus.

The paragraph beginning on page 129, line 9, has been amended as follows:

As indicated above in the fingerprinting and mapping embodiments, the present invention is also useful [to define] in defining specific stages in the temporal sequence of cells, e.g., development, and the resulting tissues within an organism. For example, the developmental stage of a cell, or population of cells, can be dependent upon the expression of particular messenger RNAs or cellular antigens. The screening procedures provided allow for high resolution definition of new classes of cells. In addition, the temporal development of particular cells will be characterized by the presence or expression of various mRNAs. Means to simultaneously screen a plurality or very large number of different sequences [as] are provided. The combination of different markers made available dramatically increases the ability to

distinguish fairly closely related cell types. Other markers may be combined with markers and methods made available herein to define new classifications of biological samples, e.g., based upon new combinations of markers.

The paragraph beginning on page 130, line 5, has been amended as follows:

In an alternative embodiment, a plurality of antigens or specific binding proteins attached to the substrate may be used to define particular cell types. For example, subclasses of T-cells are defined, in part, [upon] by the combination of expressed cell surface antigens. The present invention allows for the simultaneous screening of a large plurality of different antigens together. Thus, higher resolution classification of different T-cell subclasses becomes possible and, with the definitions and functional differences which correlate with those antigenic or other parameters, the ability to purify those cell types becomes available. This is applicable not only to T-cells, lymphocyte cells, or even to freely circulating cells. Many of the cells for which this would be most useful will be immobile cells found in particular tissues or organs. Tumor cells will be diagnosed or detected using these fingerprinting techniques. Coupled with a temporal change in structure, developmental classes may also be selected and defined using these technologies. The present invention also provides the ability not only to define new classes of cells based upon functional or structural differences, but it also provides the ability to select or purify populations of cells which share these particular properties. Standard cell sorting procedures using antibody markers may be used to detect extracellular features. Intracellular features would also be [amendable] detectable by introducing the label reagents into the cell. In particular, antisense DNA or RNA molecules may be introduced into a cell to detect RNA sequences therein. See, e.g., Weintraub (1990) Scientific American 262:40-46.

The paragraph beginning on page 130, line 35, has been amended as follows:

In an additional embodiment, the present invention also allows for the high resolution correlation of medical conditions with various different markers. For example, the [present] presently available technology, when applied to amniocentesis or other genetic screening methods, typically [screen] screens for tens of different markers at most. The present invention allows simultaneous screening for tens, hundreds, thousands, tens of thousands, hundreds of thousands, and even millions of different genetic sequences. Thus, applying the fingerprinting methods of the present invention to a sufficiently large population allows detailed statistical analysis to be made, thereby correlating particular medical conditions with particular

markers, typically antigenic or genetic. Tumor specific antigens will be identified using the present invention.

The paragraph beginning on page 131, line 12, has been amended as follows:

Various medical conditions may be correlated against an enormous data base of the sequences within an individual. Genetic propensities and correlations then become available and high resolution genetic predictability and correlation become much more easily performed. With the enormous data base, the reliability of the predictions [alsois] is also better tested. Particular markers which are partially diagnostic of particular medical conditions or medical susceptibilities will be identified and provide direction in further studies and more careful analysis of the markers involved. Of course, as indicated above in the sequencing embodiment, the present invention will find much use in intense sequencing projects. For example, sequencing of the entire human genome in the human genome project will be greatly simplified and enabled by the present invention.

The paragraph beginning on page 131, line 29, has been amended as follows:

The substrate is provided with a pattern of specific reagents which are positionally localized on the surface of the substrate. This matrix of positions is defined by the automated system which produces the substrate. The instrument will typically be one similar to that described in [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854, and [U.S.S.N. [__/___,___] U.S. Patent No. 5,489,678[, attorney docket number 11509-28 (automated VLSIPS)]. The instrumentation described therein is directly applicable to the applications used here. In particular, the apparatus comprises a substrate, typically a silicon containing substrate, on which positions on the surface may be defined by a coordinate system of positions. These positions can be individually addressed or detected by the VLSIPS™ Technology apparatus.

The paragraph beginning on page 132, line 4, has been amended as follows:

Typically, the VLSIPS™ Technology apparatus uses optical methods used in semiconductor fabrication applications. In this way, masks may be used to photo-activate positions for attachment or synthesis of specific sequences on the substrate. These manipulations may be automated by the types of apparatus described in [U.S.S.N. 07/462,492 (VLSIPS CIP)]

Pirrung et al. (1992) U.S. Pat. No. 5,143,854 and [U.S.S.N. [__/___,___] U.S. Patent No. 5,489,678. [, attorney docket number 11509-28 (automated VLSIPS).]

The paragraph beginning on page 133, line 17, has been amended as follows:

In a preferred embodiment, the protecting groups will be photoactivatable. The properties and uses of photoreactive protecting compounds have been reviewed. See, McCray et al., Ann. Rev. of Biophys. and Biophys. Chem. (1989) 18:239-270, which is incorporated herein by reference. Preferably, the photosensitive protecting groups will be removable by radiation in the ultraviolet (UV) or visible portion of the electromagnetic spectrum. More preferably, the protecting groups will be removable by radiation in the near UV or visible portion of the spectrum. In some embodiments, however, activation may be performed by other methods such as localized heating, electron beam lithography, laser pumping, oxidation or reduction with microelectrodes, and the like. Sulfonyl compounds are suitable reactive groups for electron beam lithography. Oxidative or reductive removal is accomplished by exposure of the protecting group to an electric current source, preferably using microelectrodes directed to the predefined regions of the surface which are desired for activation. A more detailed description of these protective groups is provided in [U.S.S.N. __/___,___]U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (automated VLSIPS),] which is hereby incorporated herein by reference.

The paragraph beginning on page 134, line 10, has been amended as follows:

One method to control the derivatization density is to highly derivatize the substrate with photochemical groups at high density. The substrate is then photolyzed for various predetermined times, which photoactivate the groups at a measurable rate, and react [then] them with a capping reagent. By this method, the density of linker groups may be modulated by using a desired time and intensity of photoactivation.

The paragraph beginning on page 134, line 17, has been amended as follows:

In many applications, the number of different sequences which may be provided may be limited by the density and the size of the substrate on which the matrix pattern is generated. In situations where the density is insufficiently high to allow the screening of the desired number of sequences, multiple substrates may be used to increase the number of sequences tested. Thus, the number of sequences tested may be increased by using a plurality of

different substrates. Because the VLSIPS™ Technology apparatus is almost fully automated, increasing the number of substrates does not lead to a significant increase in the number of manipulations which must be performed by humans. This again leads to greater reproducibility and speed in the handling of these multiple substrates.

The paragraph beginning on page 134, line 33, has been amended as follows:

The concept of using VLSIPS™ Technology generally allows a pattern or a matrix of reagents to be generated. The procedure for making the pattern is performed by any of a number of different methods. An apparatus and instrumentation useful for generating a high density VLSIPS™ Technology substrate is described in detail in [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 and [U.S.S.N. __/___,___] U.S. Patent No. 5,489,678.[, attorney docket number 11509-28 (automated VLSIPS).]

The paragraph beginning on page 135, line 5, has been amended as follows:

The details of the binary masking are described in an accompanying application filed simultaneously with this, U.S. Patent No. 5,489,678 [U.S.S.N. __/___,___ attorney docket number 11509-28 (automated VLSIPS)] whose specification is incorporated herein by reference.

The paragraph beginning on page 135, line 21, has been amended as follows:

The synthetic methods in making a substrate are described in the parent application, [U.S.S.N. 07/492,462] which issued as U.S. Pat. No. 5,143,854. The construction of the matrix pattern on the substrate will typically be generated by the use of photo-sensitive reagents. By use of photo-lithographic optical methods, particular segments of the substrate can be irradiated with light to activate or deactivate blocking agents, e.g., to protect or deprotect particular chemical groups. By an appropriate sequence of photo-exposure steps at appropriate times with appropriate masks and with appropriate reagents, the substrates can have known polymers synthesized at positionally defined regions on the substrate. Methods for synthesizing various substrates are described in [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 and U.S. Patent No. 5,489,678[U.S.S.N. __/___,___, attorney docket number 11509-28 (automated VLSIPS)]. By a sequential series of these photo-exposure and reaction manipulations, a defined matrix pattern of known sequences may be generated, and is typically referred to as a VLSIPS™ Technology substrate. In the nucleic acid synthesis

embodiment, nucleosides used in the synthesis of DNA by photolytic methods will typically be one of the two forms shown below:

The paragraph beginning on page 136, line 37, has been amended as follows:

In I, the photolabile group at the 5' position is abbreviated NV (nitroveratryl) and in II, the group is abbreviated NVOC (nitroveratryl oxycarbonyl). Although not shown in Fig. C the bases (adenine, cytosine, and guanine) contain exocyclic NH_2 groups which must be protected during DNA synthesis. Thymine contains no exocyclic NH_2 and therefore requires no protection. The standard protecting groups for these [anaines] amines are shown below:

The paragraph beginning on page 138, line 23, has been amended as follows:

Nucleosides used as 5'-OH probes, useful in verifying correct VLSIPS™ Technology synthetic function, [have been] include, for example the following:

The paragraph beginning on page 140, line 12, has been amended as follows:

If the surface is hydroxy functionalized, a phosphate bond is made (see example below):

The paragraph beginning on page 140, line 33, has been amended as follows:

One of the most common 3'-O-protecting [group] groups is the ester, in particular the acetate:

The paragraph beginning on page 141, line 3, has been amended as follows:

Another group used most often is the silyl ether[.]:

The paragraph beginning on page 141, line 16, has been amended as follows:

[Related to] With respect to photodeprotection, the nitroveratryl group could also be used to protect the 3'-position.

The paragraph beginning on page 141, line 29, has been amended as follows:

A variety of ethers can also be used in the protection of the 3'-O-position[.]:

The paragraph beginning on page 142, line 3, has been amended as follows:

Note that corresponding linkages and photoblocked amino acids are described in detail in U.S. Patent No. 5,489,678[U.S.S.N. __/___,___, attorney docket number 11509-28], which is hereby incorporated herein by reference.

The paragraph beginning on page 142, line 16, has been amended as follows:

As an alternative method of generating a matrix pattern on a substrate, preformed polymers may be individually attached at particular sites on the substrate. This may be performed by individually attaching reagents one at a time to specific positions on the matrix, a process which may be automated. See, e.g., U.S.S.N. 07/435,316[(caged biotin parent), and U.S.S.N. 07/612,671 (caged biotin CIP)] , from which CIP U.S.S.N. 07/612,671 issued as U.S. Pat. No. 5,252,743. Another way of generating a positionally defined matrix pattern on a substrate is to have individually specific reagents which interact with each specific position on the substrate. For example, oligonucleotides may be synthesized at defined locations on the substrate. Then the substrate would have on its surface a plurality of regions having homogeneous oligonucleotides attached at each position.

The paragraph beginning on page 142, line 30, has been amended as follows:

In particular, at least four different substrate preparation procedures are available for treating a substrate surface. They are the standard VLSIPS™ Technology method, polymeric substrates, Durapore™, and synthetic beads or fibers. The treatment labeled "standard VLSIPS™ Technology " method is described in U.S Patent No. 5,489,678[U.S.S.N. __/___,___], [attorney docket number 11509-28 (automated VLSIPS),] and involves applying amino-propyltriethoxysilane to a glass surface.

The paragraph beginning on page 143, line 1, has been amended as follows:

The polymeric substrate approach involves either of two ways of generating a polymeric substrate. The first uses a high concentration of aminopropyltriethoxysilane (2-20%) in an aqueous ethanol solution (95%). This allows the silane compound to polymerize both in solution and on the substrate surface, which provides a high density of amines on the surface of the glass. This density is contrasted with the standard VLSIPS™ Technology method. This polymeric method allows for the deposition on the substrate surface of a monolayer due to the anhydrous method used with the aforementioned silane.

The paragraph beginning on page 143, line 35, has been amended as follows:

Here the building up of, e.g., a DNA oligomer, can be started immediately since phosphate bonds to the surface can be accomplished in the first step with no need for modification. A nucleotide dimer (5'-C-T-3') has been successfully made on this substrate[in our labs].

The paragraph beginning on page 144, line 3, has been amended as follows:

The fourth method utilizes synthetic beads or fibers. This would use another substrate, such as a teflon copolymer graft bead or fiber, which is covalently coated with an organic layer (hydrophilic) terminating in hydroxyl sites (commercially available from Molecular [Brosystems] Biosystems, Inc.) This would offer the same advantage as the Durapore™ membrane, allowing for immediate phosphate linkages, but would give additional contour by the 3-dimensional growth of oligomers.

The paragraph beginning on page 145, line 25, has been amended as follows:

An alternative method of attaching reagents in a positionally defined matrix pattern is to use a caged biotin system. See [U.S.S.N. 07/612,671 (caged biotin CIP)] Barrett et al. (1993) U.S. Pat. No. 5,252,743, which is hereby incorporated herein by reference, for additional details on the chemistry and application of caged biotin embodiments. In short, the caged biotin has a photosensitive blocking moiety which prevents the combination of avidin to biotin. At positions where the photo-lithographic process has removed the blocking group, high affinity biotin sites are generated. Thus, by a sequential series of photolithographic deblocking steps interspersed with exposure of those regions to appropriate biotin containing reagents, only those locations where the deblocking takes place will form an avidin-biotin interaction. Because the avidin-biotin binding is very tight, this will usually be virtually irreversible binding.

The paragraph beginning on page 146, line 31, has been amended as follows:

As discussed previously in the VLSIPS™ Technology parent applications, the VLSIPS™ Technology substrates may be used for screening for specific interactions with sequence specific targets or probes.

The paragraph beginning on page 148, line 1, has been amended as follows:

In another embodiment, a fingerprinting-like procedure may be used for classifying cell types by analyzing a pattern of specific nucleic acids present in the cell. A series of antibodies may be used to identify cell markers, e.g., proteins, usually on the cell surface, but intracellular markers may also be used. Antigens which are extracellularly expressed are preferred so cell lysis is unnecessary in the screening, but intracellular markers may also be useful. The markers will usually be proteins, but may be nucleic acids, lipids, metabolites, carbohydrates, or other cellular components. See, e.g., Winkelgren, I. (1990) Science News 136:234-237, which indicates extracellular DNA may be common, and suggesting that such might be characteristic of cell types, stage, or physiology. This may also be useful in defining the temporal stage of development of cells, e.g., stem cells or other cells which undergo temporal changes in development. For example, the stage of a cell, or group of cells, may be tested or defined by isolating a sample of mRNA from the population and testing to see what sequences are present in messenger populations. Direct samples, or amplified samples, may be used. Where particular mRNA or other nucleic acid sequences may be characteristic of or shown to be characteristic of particular developmental stages, physiological states, or other conditions, this fingerprinting method may define them. Similar sorts of fingerprinting may be used for determining T-cell classes or perhaps even to generate classification schemes for such proteins as major histocompatibility complex antigens. Thus, the ability to make these substrates allows both the generation of reagents which will be used for defining subclasses or classes of cells or other biological materials, but also provides the mechanisms for selecting those cells which may be found in defined population groups.

The paragraph beginning on page 148, line 33, has been amended as follows:

[Cell] In addition to cell classification defined by such a combination of properties, typically expression of extracellular antigens, the present invention also provides the means for isolating homogeneous population of cells. Once the antigenic determinants which define a cell class have been identified, these antigens may be used in a sequential selection process to isolate only those cells which exhibit the combination of defining structural properties.

The paragraph beginning on page 149, line 3, has been amended as follows:

The present invention may also be used for mapping sequences within a larger segment. This may be performed by at least two methods, particularly in reference to nucleic acids. Often, enormous segments of DNA are subcloned into a large plurality of subsequences.

Ordering these subsequences may be important in determining the overlaps of sequences upon nucleotide determinations. Mapping may be performed by immobilizing particularly large segments onto a matrix using the VLSIPS™ technology. Alternatively, sequences may be ordered by virtue of subsequences shared by overlapping segments. See, e.g., Craig et al. (1990) Nuc. Acids Res. 18:2653-2660; Michiels et al. (1987) CABIOS 3:203-210; and Olson et al. (1986) Proc. Natl. Acad. Sci. USA 83:7826-7830.

The paragraph beginning on page 149, line 18, has been amended as follows:

The extent of specific interaction between reagents immobilized to the VLSIPS™ Technology substrate and another sequence specific reagent may be modified by the conditions of the interaction. Sequencing embodiments typically require high fidelity hybridization and the ability to discriminate perfect matching from imperfect matching. Fingerprinting and mapping embodiments may be performed using less stringent conditions, depending upon the circumstances.

The paragraph beginning on page 151, line 15, has been amended as follows:

The rate of hybridization can also be affected by the inclusion of particular hybridization accelerators. These hybridization accelerators include the volume exclusion agents characterized by dextran sulfate, or polyethylene glycol (PEG). Dextran sulfate is typically included at a concentration of between 1% and 40% by weight. The actual concentration selected depends upon the application, but typically a faster hybridization is desired in which the concentration is optimized for the system in question. Dextran sulfate is often included at a concentration of between 0.5% and 2% by weight or dextran sulfate at a concentration between about 0.5% and 5%. Alternatively, proteins which accelerate hybridization may be added, e.g., the recA protein found in E. coli[)] or other homologous proteins.

The paragraph beginning on page 153, line 25, has been amended as follows:

Suitable chromogens will include molecules and compounds which absorb light in a distinctive range of wavelengths so that a color may be observed, or emit light when irradiated with radiation of a particular wave length or wave length range, e.g., fluorescers. Biliproteins, e.g., [ficoerythrin] phycoerythrin, may also serve as labels.

The paragraph beginning on page 153, line 31, has been amended as follows:

A wide variety of suitable dyes are available, being [primary] primarily chosen to provide an intense color with minimal absorption by their surroundings. Illustrative dye types include quinoline dyes, triarylmethane dyes, acridine dyes, alizarine dyes, phthaleins, insect dyes, azo dyes, anthraquinoid dyes, cyanine dyes, phenazathionium dyes, and phenazoxonium dyes.

The paragraph beginning on page 156, line 7, has been amended as follows:

With the automated detection apparatus, the correlation of specific positional labeling is converted to the presence on the target of sequences for which the reagents have specificity of interaction. Thus, the positional information is directly converted to a database indicating what sequence interactions have occurred. For example, in a nucleic acid hybridization application, the sequences which have interacted between the substrate matrix and the target molecule can be directly listed from the positional information. The detection system used is described in[U.S.S.N. 07/649,642 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854; and [U.S.S.N. __/___,___] U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (automated VLSIPS)]. Although the detection described therein is a fluorescence detector, the detector may be replaced by a spectroscopic or other detector. The scanning system may make use of a moving detector relative to a fixed substrate, a fixed detector with a moving substrate, or a combination. Alternatively, mirrors or other apparatus can be used to transfer the signal directly to the detector. See, e.g, [U.S.S.N. __/___,___, attorney docket number 11509-28 (automated VLSIPS),] U.S. Patent No. 5,489,678, which is hereby incorporated herein by reference.

The paragraph beginning on page 157, line 19, has been amended as follows:

More sophisticated signal processing techniques can be applied to the initial determination of whether a positive signal exists or not. See, e.g., U.S. Patent No. 5,489,678[S.N. __/___,___, attorney docket number 11509-28 (automated VLSIPS)].

The paragraph beginning on page 157, line 23, has been amended as follows:

From a listing of those sequences which interact, data analysis may be performed on a series of sequences. For example, in a nucleic acid sequence application, each of the sequences may be analyzed for their overlap regions and the original target sequence may be reconstructed from the collection of specific subsequences obtained therein. Other sorts of analyses for different applications may also be performed, and because the scanning system

directly interfaces with a computer the information need not be transferred manually. This provides for the ability to handle large amounts of data with very little human intervention. This, of course, provides significant advantages over manual manipulations. Increased throughput and reproducibility is thereby provided by the automation of a vast majority of steps in any of these applications.

The paragraph beginning on page 158, line 3, has been amended as follows:

Data analysis will typically involve aligning the proper sequences with their overlaps to determine the target sequence. Although the target "sequence" may not specifically correspond to any specific molecule, especially where the target sequence is broken and fragmented[up] in the sequencing process, the sequence corresponds to a contiguous sequence of the subfragments.

The paragraph beginning on page 161, line 29, has been amended as follows:

As indicated above, although the VLSIPS™ Technology may be applied to sequencing embodiments, it is often useful to integrate other concepts to [simply] simplify the sequencing. For example, nucleic acids may be easily sequenced by careful selection of the vectors and hosts used for amplifying and generating the specific target sequences. For example, it may be desired to use specific vectors which have been designed to interact most efficiently with the VLSIPS™ Technology substrate. This is also important in fingerprinting and mapping strategies. For example, vectors may be carefully selected having particular complementary sequences which are designed to attach to a genetic or specific oligomer on the substrate. This is also applicable to situations where it is desired to target particular sequences to specific locations on the matrix.

The paragraph beginning on page 162, line 5, has been amended as follows:

In one embodiment, unnatural oligomers may be used to target natural probes to specific locations on the VLSIPS™ Technology substrate. In addition, particular probes may be generated for the mapping embodiment which are designed to have specific combinations of characteristics. For example, the construction of a mapping substrate may depend upon use of another automated apparatus which takes clones isolated from a chromosome walk and attaches them individually or in bulk to the VLSIPS™ Technology substrate.

The paragraph beginning on page 162, line 14, has been amended as follows:

In another embodiment, a variety of specific vectors having known and particular "targeting" sequences adjacent to the cloning sites may be individually used to clone a selected probe, and the isolated probe will then be targetable to a site on the VLSIPS™ Technology substrate with a sequence complementary to the "target" sequence.

The paragraph beginning on page 163, line 7, has been amended as follows:

Also as indicated above, various means for constructing an appropriate substrate may involve either mechanical or automated procedures. The standard VLSIPS™ Technology automated procedure involves synthesizing oligonucleotides or short polymers directly on the substrate. In various other embodiments, it is possible to attach separately synthesized reagents onto the matrix in an ordered array. Other circumstances may lend themselves to transfer a pattern from a petri plate onto a solid substrate. Also, there are methods for site specifically directing collections of reagents to specific locations using unnatural nucleotides or equivalent sorts of targeting molecules.

The paragraph beginning on page 163, line 19, has been amended as follows:

While a brute force manual transfer process may be utilized sequentially for attaching various samples to successive positions, instrumentation for automating such procedures may also be devised. The automated system for performing such would preferably be relatively easily designed and conceptually easily understood.

The paragraph beginning on page 164, line 6, has been amended as follows:

Of course, sequencing[of] can be very important in many different sorts of environments. For example, it will be useful in determining the genetic sequence of particular markers in various individuals. In addition, polymers may be used as markers or for information containing molecules to encode information. For example, a short polynucleotide sequence may be included in large bulk production samples indicating the manufacturer, date, and location of manufacture of a product. For example, various drugs may be encoded with this information with a small number of molecules in a batch. For example, a pill may have somewhere from 10 to 100 to 1,000 or more very short and small molecules encoding this information. When necessary, this information may be decoded from a sample of the material using a polymerase chain reaction (PCR) or other amplification method. This encoding system may be used to

provide the origin of large bulky samples without significantly affecting the properties of those samples. For example, chemical samples may also be encoded by this method thereby providing means for identifying the source and manufacturing details of lots. The origin of bulk hydrocarbon samples may be encoded. Production lots of organic compounds such as benzene or plastics may be encoded with a short molecule polymer. Food stuffs may also be encoded using similar marking molecules. Even toxic waste samples can be encoded determining the source or origin. In this way, proper disposal can be traced or more easily enforced.

The paragraph beginning on page 165, line 12, has been amended as follows:

As indicated above, fingerprinting technology may also be used for data encryption. Moreover, fingerprinting allows for significant identification of particular individuals. Where the fingerprinting technology is standardized, and used for identification of large numbers of people, related equipment and peripheral processing will be developed to accompany the underlying technology. For example, specific equipment may be developed for automatically taking a biological sample and generating or amplifying the information molecules within the sample to be used in fingerprinting analysis. Moreover, the fingerprinting substrate may be mass produced using particular types of automatic equipment. Synthetic equipment may produce the entire matrix simultaneously by stepwise synthetic methods as provided by the VLSIPS™ technology. The attachment of specific probes onto a substrate may also be automated, e.g., making use of the caged biotin technology. See, e.g., [U.S.S.N. 07/612,671 (caged biotin CIP)] Barrett et al. (1993) U.S. Pat. no. 5,252,743. As indicated above, there are automated methods for actually generating the matrix and substrate with distinct sequence reagents positionally located at each of the matrix positions. Where such reagents are, e.g., unnatural amino acids, a targeting function may be utilized which does not interfere with [a]a natural nucleotide functionality.

The paragraph beginning on page 166, line 21, has been amended as follows:

In particular industries, the VLSIPS™ Technology sequencing, fingerprinting, or mapping technology will be particularly appropriate. As mentioned above, agricultural livestock suppliers may be able to encode and determine whether their particular strains are being used by others. By incorporating particular markers into their genetic stocks, the markers will indicate origin of genetic material. This is applicable to seed producers, livestock producers, and other suppliers of medical or agricultural biological materials.

The paragraph beginning on page 167, line 3, has been amended as follows:

As indicated above, polymers may be used to encode important information on source and batch and supplier. This is described in greater detail, e.g., "Applications of PCR to industrial problems," (1990) in Chemical and Engineering News 68:145, which is hereby incorporated herein by reference. In fact, the synthetic method can be applied to the storage of enormous amounts of information. Small substrates may encode enormous amounts of information, and its recovery will make use of the inherent replication capacity. For example, on regions of 10 μm x 10 μm, 1 cm^2 has 10^6 regions. [An] In theory, the entire human genome could be attached in 1000 nucleotide segments on a 3 cm^2 surface. Genomes of endangered species may be stored on these substrates.

The paragraph beginning on page 167, line 35, has been amended as follows:

Modifications of the fingerprint embodiments may be used to diagnose the condition of the organism. For example, a blood sample is presently used for diagnosing any of a number of different physiological conditions. A multi-dimensional fingerprinting method made available by the present invention could become a routine means for diagnosing an enormous number of physiological features simultaneously. This may revolutionize the practice of medicine in providing information on an enormous number of parameters together at one time. In another way, the genetic predisposition may also revolutionize the practice of medicine providing a physician with the ability to predict the likelihood of particular medical conditions arising at any particular moment. It also provides the ability to apply [preventative] preventive medicine.

The paragraph beginning on page 170, line 36, has been amended as follows:

Relevant applications whose techniques are incorporated herein by reference are Pirrung, et al., U.S.S.N. 07/362,901 [(VLSIPS parent), filed June 7, 1989, now abandoned;] , from which CIP 07/492,462 issued as U.S. Patent No. 5,143,854, [Pirrung et al., U.S.S.N. 07/492,462 (VLSIPS CIP) 1992 U.S. Pat. No. 5,143,854, filed March 7, 1990; Barrett, et al.,] U.S.S.N. 07/435,316, from which CIP 07/612,671 issued as U.S. Patent No. 5,252,743,[(caged biotin) filed November 13, 1989; Barrett, et al.,U.S.S.N. 07/612,671 (caged biotin CIP), filed November 13, 1990]and U.S. Patent No. 5,489,678.[; and commonly assigned and simultaneously filed applications U.S.S.N. __/___,___, attorney docket number 11509-28

(automated VLSIPS) and U.S.S.N. __/___,___, attorney docket number 11509-26 (sequencing by synthesis).]

The paragraph beginning on page 171, line 33, has been amended as follows:

In order to determine the time for photolysis of [5'-O-nitrovertryl] 5'-O-nitroveratryl thymidine to thymidine a 100 μM solution of NV-Thym-OH ([5'-O-nitrovertryl] 5'-O-nitroveratryl thymidine) in dioxane was made and ~200 μl aliquots were irradiated (in a quartz cuvette 1 cm x 2 mm) at 362.3 nm for 20 sec, 40 sec, 60 sec, 2 min, 5 min, 10 min, 15 min, and 20 min. The resulting irradiated mixtures were then analyzed by HPLC using a Varian MicroPak SP column (C_{18} analytical) at a flow rate of 1 ml/min and a solvent system of 40% CH3CN and 60% water. Thymidine has a retention time of 1.2 min and NVO-Thym-OH has a retention time of 2.1 min. It was seen that after 10 min of exposure the deprotection was complete.

The paragraph beginning on page 172, line 31, has been amended as follows:

To an aminopropylated glass slide (standard VLSIPS™ Technology) was added a mixture of the following:

12.2 mg of NVO-Thym-CO_2H (IX)
3.4 mg of HOBT (N-hydroxybenztriazl)
8.8 μl DIEA (Diisopropylethylamine)
11.1 mg BOP reagent
2.5 ml DMF

The paragraph beginning on page 174, line 3, has been amended as follows:

An aminopropyl glass slide, was soaked in a solution of ethylene oxide (20% in DMF) to generate a hydroxylated surface. The slide was added to a mixture of the following:

32 mg of NVO-T-OCED (X)
11 mg of tetrazole
0.5 mg of anhydrous CH_3CN

The paragraph beginning on page 174, line 9, has been amended as follows:

After 8 min the plate was then rinsed with acetonitrile, then oxidized with I_2/H_2O/THF/lutidine for 1 min, washed and dried. The slide was then exposed to a 1:3 mixture of acetic anhydride:pyridine for 1 h, then washed and dried. The substrate was [a] then

photolyzed in dioxane at 362 nm at 14 mW/cm^2 for 10 min using a 500μm checkerboard mask, dried, and then treated with a mixture of the following:

65 mg of biotin modified C (IV)

11 mg of tetrazole

0.5 ml anhydrous CH_3CN

The paragraph beginning on page 177, line 27, has been amended as follows:

The above example has illustrated the method by way of a manual example. It will of course be appreciated that automated or semi-automated methods could be used. The substrate would be mounted in a flow cell for automated addition and removal of reagents, to minimize the volume of reagents needed, and to more carefully control reaction conditions. Successive masks will be applicable manually or automatically. See, e.g., [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 and [U.S.S.N. [__/___,___] U.S. Patent No. 5,489,678[, attorney docket number 11509-28 (automated VLSIPS)].

The paragraph beginning on page 178, line 20, has been amended as follows:

The described technique for making dimers of A, C, G, and T may be further extended to make longer oligonucleotides. The automated system described, e.g., in[U.S.S.N 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854, and [U.S.S.N. __/___,___] U.S. Patent No. 5,489,678, [attorney docket number 11509-28 (automated VLSIPS),] can be adapted to make all possible 10-mers composed of the 4 nucleotides A, C, G, and T. The photosensitive, blocked nucleotide analogues have been described above, and would be readily adaptable to longer oligonucleotides.

The paragraph beginning on page 180, line 18, has been amended as follows:

As indicated above, the detection of specific interactions may be performed by detecting the positions where the labeled target sequences are attached. Where the label is a fluorescent label, the apparatus described, e.g., in[U.S.S.N. 07/492,462 (VLSIPS CIP);] Pirrung et al. (1992) U.S. Pat. No. 5,143,854, [and U.S.S.N. __/___,___ attorney docket number 11509-28,] may be advantageously applied. In particular, the synthetic processes described above will result in a matrix pattern of specific sequences attached to the substrate, and a known pattern of interactions can be converted to corresponding sequences.

The paragraph beginning on page 181, line 9, has been amended as follows:

The description of the preparation of short peptides on a substrate incorporates by reference sections in [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854, and described below.

The paragraph beginning on page 181, line 18, has been amended as follows:

The aminated surface of the slide is exposed to about 500 µl of, e.g., a 30 millimolar (mM) solution of NVOC-GABA (gamma amino butyric acid) NHS (N-hydroxysuccinimide) in DMF for attachment of a NVOC-GABA to each of the amino groups. The surface is washed with, for example, DMF, methylene chloride, and ethanol. See [U.S.S.N. [__/___,___,attorney docket number 11509-28,] U.S. Patent No. 5,489,678 for details on amino acid chemistry.

The paragraph beginning on page 181, line 36, has been amended as follows:

See [U.S.S.N. 07/492,462(VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 which describes the preparation of glycine and phenylalanine trimers. The technique is similar to the method described above for making triners of C and T, but substituting photosensitive blocked glycine for the C derivative and photosensitive blocked phenylalamine for the T derivative.

The paragraph beginning on page 182, line 7, has been amended as follows:

In synthesizing the dimer of an aminopropyl group and a fluorescent group, a functionalized [durapore]Durapore™ membrane was used as a substrate. The Durapore™ membrane was a polyvinylidine difluoride with aminopropyl groups. The aminopropyl groups were protected with the DDZ group by reaction of the carbonyl chloride with the amino groups, a reaction readily known to those of skill in the art. The surface bearing these groups was placed in a solution of THF and contacted with a mask bearing a checkerboard pattern of 1 mm opaque and transparent regions. The mask was exposed to ultraviolet light having a wavelength down to at least about 280 nm for about 5 minutes at ambient temperature, although a wide range of exposure times and temperatures may be appropriate in various embodiments of the invention. For example, in one embodiment, an exposure time of between about 1 and 5000 seconds may be used at process temperatures of between -70 and +50°C.

The paragraph beginning on page 183, line 8, has been amended as follows:

One of the beads was placed in the illumination field on the scan stage in a field of a laser spot which was initially shuttered. After being positioned in the illumination field, the photon detection equipment was turned on. The laser beam was unblocked and it interacted with the particle bead, which then fluoresced. Fluorescence curves of beads impregnated with 7,000 and 29,000 fluorescein molecules, are shown in Figs. 11A and 11B, respectively of [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854. On each curve, traces for beads without fluorescein molecules are also shown. These experiments were performed with 488 nm excitation, with 100 μW of laser power. The light was focused through a 40 power 0.75 NA objective.

The paragraph beginning on page 184, line 13, has been amended as follows:

Through the use of curves similar to those shown in Fig. 11 of [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854, and by integrating the fluorescent counts under the exponentially decaying signal, the number of free amino groups on the surface after derivitization was determined. It was determined that slides with labeling densities of 1 fluorescein per 10^3x10^3 to ~2x2 nm could be reproducibly made as the concentration of aminopropyltriethoxysilane varied from 10^{-5}% to 10^{-1}%.

The paragraph beginning on page 184, line 22, has been amended as follows:

7. removal of [NOVC] NVOC and attachment of a fluorescent marker

The paragraph beginning on page 184, line 29, has been amended as follows:

Fig. 12A of [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854 illustrates the slide which was not exposed to light, but which was exposed to FITC. The units of the x axis are time and the units of the y axis are counts. The trace contains a certain amount of background fluorescence. The duplicate slide was exposed to 350 nm broadband illumination for about 1 minute (12 mW/cm^2, ~350 nm illumination), washed and reacted with FITC. A large increase in the level of fluorescence is observed, which indicates photolysis has exposed a number of amino groups on the surface of the slides for attachment of a fluorescent marker.

The paragraph beginning on page 186, line 1, has been amended as follows:

The results of this experiment were plotted as fluorescence intensity as a function of position. This image was taken at 10 μm steps and showed that not only can deprotection be carried out in a well defined pattern, but also that (1) the method provided for successful coupling of peptides to the surface of the substrate, (2) the surface of a bound peptide was available for binding with an antibody, and (3) [that] the detection apparatus capabilities were sufficient to detect binding of a receptor. Moreover, the Herz antibody is a sequence specific reagent which may be used advantageously as a sequence specific recognition reagent. It may be used, if specificity is high, for sequencing purposes, and, at least, for fingerprinting and mapping uses.

The paragraph beginning on page 186, line 21, has been amended as follows:

A slide is derivatized with the aminopropyl group, protected in this case with t-BOC (t-butoxycarbonyl). The slide was treated with TFA to remove the t-BOC protecting group. E-aminocaproic acid, which was t-BOC protected at its amino group, was then coupled onto the aminopropyl groups. The aminocaproic acid serves as a spacer between the aminopropyl group and the peptide to be synthesized. The amino end of the spacer was deprotected and coupled to NVOC-leucine. The entire slide was then illuminated with 12 mW of 325 nm broadband illumination. The slide was then coupled with NVOC-phenylalanine and washed. The entire slide was again illuminated, then coupled to NVOC-glycine and washed. The slide was again illuminated and coupled to NVOC-glycine to form the sequence shown in the last portion of Fig. 13A of [U.S.S.N. 07/492,462 (VLSIPS CIP)] Pirrung et al. (1992) U.S. Pat. No. 5,143,854.

The paragraph beginning on page 190, line 19, has been amended as follows:

According to an alternative embodiment of the invention, the methods provide for attaching to the surface a caged binding member which, in its caged form, has a relatively low affinity for other potentially binding species, such as receptors and specific binding substances. Such techniques are more fully described in copending application Serial No. 404,920, filed September 8, 1989, and incorporated herein by reference for all purposes. See also U.S.S.N. 07/435,316, [(caged biotin parent) and] from which CIP U.S.S.N. 07/612,671 [(caged biotin CIP)] issued as U.S. Pat. No. 5,252,743, each of which is hereby incorporated herein by reference.

The paragraph beginning on page 191, line 32, has been amended as follows:

Preferably, the caged binding member attached to the solid substrate will be a photoactivatable biotin complex, i.e., a biotin molecule that has been chemically modified with photoactivatable protecting groups so that it has a significantly reduced binding affinity for avidin or avidin analogs than does natural biotin. In a preferred embodiment, the protecting groups localized in a predefined region of the surface will be removed upon application of a suitable source of radiation to give binding members, that [are] is biotin or a functionally analogous compound having substantially the same binding affinity for avidin or avidin analogs as does biotin.

The paragraph beginning on page 192, line 29, has been amended as follows:

Polynucleotide fingerprinting may use reagents similar to those described above for probing a sequence for the presence of specific subsequences found therein. Typically, the subsequences used for fingerprinting will be longer than the sequences used in oligonucleotide sequencing. In particular, specific long segments may be used to determine the similarity of different samples of nucleic acids. They may also be used to fingerprint whether specific combinations of information are provided therein. Particular probe sequences are selected and attached in a positional manner to a substrate. The means for attachment may be either using a caged biotin method described, e.g., in [U.S.S.N. 07/612,671 (caged biotin CIP)] Barrett et al. (1993) U.S. Pat. No. 5,252,743, or by another method using targeting molecules. For example, a short polypeptide of specific sequence may be attached to an oligonucleotide and targeted to specific positions on a substrate having antibodies attached thereto, the antibodies exhibiting specificity for binding to those short peptide sequences. In another embodiment, an unnatural nucleotide or similar complementary binding molecule may be attached to the fingerprinting probe and the probe thereby directed towards complementary sequences on a VLSIPS™ Technology substrate. Typically, unnatural nucleotides would be preferred, e.g., unnatural optical isomers, which would not interfere with natural nucleotide interactions.

The paragraph beginning on page 197, line 1, has been amended as follows:

In one embodiment, the two may be combined in a single incubation step. A particular incubation condition may be found which is compatible with both hybridization recognition and non-hybridization recognition molecules. Thus, e.g., an incubation condition may be selected which allows both specificity of antibody binding and specificity of nucleic acid hybridization. This allows simultaneous performance of both types of interactions on a single

matrix. Again, where developmental mRNA patterns are correlated with structural features, or with probes which are able to hybridize to intracellular mRNA populations, a cell sorter may be used to sort specifically those cells having desired mRNA population patterns.

The paragraph beginning on page 198, line 30, has been amended as follows:

In addition, an allergy like test may be used to diagnose the immunological history of a particular individual. For example, by testing the circulating antibodies in a blood sample, which reflects the immunological history and memory of an individual, it may be determined what infections may not have been historically presented to the immune system. In this manner, it may be possible to specifically supplement an immune system for a short period of time with IgG fractions made up of specific types of gamma globulins. Thus, hepatitis gamma globulin injections may be better designed for a particular environment to which a person is expected to be exposed. This also provides the ability to identify genetically equivalent individuals who have immunologically different experiences. Thus, a blood sample from an individual who has a particular combination of circulating antibodies will likely be different from the combination of circulating antibodies found in a genetically similar or identical individual. This could allow for the distinction between clones of particular animals, e.g., mice, rats, or other animals.

The paragraph beginning on page 200, line 5, has been amended as follows:

As indicated above in the diagnostic tests, it is possible to identify a particular immune system within a genetically homogeneous class of organisms by virtue of [her] their immunological history. For example, a large colony of cloned mice may be distinguishable by virtue of each immunological history. For example, one mouse may have had an immunological response to exposure to antigen A to which her genetically identical sibling may have not been exposed. By virtue of this differential history, the first of the pair will likely have a high antibody titer against the antigen A whereas her genetically identical sibling will have not had a response to that antigen by virtue of never having been exposed to it. For this reason, immune systems may be identified by their immunological memories. Thus, immunological experience may also be a means for identifying a particular individual at a particular moment in her lifetime.

The paragraph beginning on page 201, line 33, has been amended as follows:

In addition, the specific probes may be positionally directed to specific locations on a substrate by targeting. For example, polypeptide specific recognition reagents may be

attached to oligonucleotide sequences which can be complementarily targeted to specific locations on a VLSIPS™ Technology substrate. Hybridization conditions, as applied for oligonucleotide probes, will be used to target the reagents to locations on a substrate having complementary oligonucleotides synthesized thereon. In another embodiment, oligonucleotide probes may be attached to specific polypeptide targeting reagents such as an antigen or antibody. These reagents can be directed towards a complementary antigen or antibody already attached to a VLSIPS™ Technology substrate.

The paragraph beginning on page 202, line 19, has been amended as follows:

Positionally defined clones may be transferred to a new substrate by either physical transfer or by synthetic means. Synthetic means may involve either a production of the probe on the substrate using the VLSIPS™ Technology synthetic methods, or may involve the attachment of a targeting sequence made by VLSIPS™ Technology synthetic methods which will target that positionally defined clone to a position on a new substrate. Both methods will provide a substrate having a number of positionally defined probes useful in mapping.